

Year ahead prediction of US landfalling hurricane numbers: the optimal combination of multiple levels of activity since 1900

Roman Binter (LSE)
 Stephen Jewson (RMS)*
 Shree Khare (RMS)
 Adam O'Shay (RMS)
 Jeremy Penzer (LSE)

February 2, 2008

Abstract

In earlier work we considered methods for predicting future levels of hurricane activity based on the assumption that historical mean activity was at one constant level from 1900 to 1994, and has been at another constant level since then. We now make this model a little more subtle, and account for the possibility of four different levels of mean hurricane activity since 1900.

1 Introduction

We want to predict the number of hurricanes that might occur during the next hurricane season, and we want to make that prediction as soon as the current season is over. We call this problem ‘year-ahead’ prediction of hurricane numbers, to distinguish it from ‘seasonal’ predictions of hurricane numbers, which are generally made from a point in time closer to the hurricane season being predicted. Seasonal predictions often use predictors in the atmosphere and ocean that correlate with hurricane activity (see, for example, Saunders and Lea (2005), Gray (1984), Gray et al. (1992), Gray (1993) and Gray (1994), Elsner and Schmertmann (1993), Landsea et al. (1999), Landsea (2000), Lehmiller et al. (1997)). These predictors are generally less useful, however, at the long lead time that we are considering, and we therefore take a rather different approach, based on the idea of using linear combinations of the historical hurricane number data as a prediction of the future number of hurricanes. The particular linear combinations we consider put different weights on different periods of historical data according to how similar those data are to the current levels of activity.

Our first attempt in this direction (described in Jewson et al. (2005)) was to assume that historical hurricane activity was at one (constant) level of mean activity from 1900 to 1994, and jumped to another (constant) level of mean activity from 1995 to 2005. We then assumed that 2006 would remain at the new level of mean activity, and predicted this new level of activity using a combination of the estimated levels of mean activity from 1900 to 1994 and 1995 to 2005. We formulated the problem mathematically as a mean-variance trade-off: the recent data is (according to our assumptions) unbiased, but too short to give accurate estimates. The earlier data is biased, but long enough to make our estimates much more accurate. An appropriately weighted combination of the data from the two periods can then give a prediction with MSE lower than the predictions from either period individually, and we derived expressions that give an estimate of the weights needed to make this combination.

Perhaps the biggest objection that one might make to this initial study (in terms of realism), is that no-one really thinks that hurricane activity was at one constant level from 1900-1994. A more common point of view is that activity levels were low in the early part of the 20th century, higher in the middle of the century, low again until the mid 1990s, and then high again. The goal of this article is to extend our analysis to account for this idea. Instead of modelling hurricane numbers in two periods of constant levels of activity, we now model it in four periods. Otherwise, our method is exactly as before: we consider how to make the combinations of the estimated levels of activity in these four periods, with the goal of minimising the MSE of predictions for next year, assuming that next year remains at the same level of

* Correspondence email: stephen.jewson@rms.com

mean activity as the current period. Within this framework we consider a number of different ways of weighting the four periods of activity.

The periods we use are:

- period 4: 1900-1942
- period 3: 1943-1964
- period 2: 1965-1994
- period 1: 1995-2005

The definitions of these periods are taken from the change-point analysis in Elsner et al. (2000), figure 2. It would be interesting to investigate how robust these periods are: however, the point of this article is to demonstrate how to combine data, given the definitions of the four periods, rather than to investigate the definitions of the periods themselves. We number the periods in reverse chronological order to emphasize that period 1 is unique in that we are assuming that 2006 will be at the same activity level as period 1. The rest of this paper is as follows. In section 2 we present the hurricane data we will use for our analysis. In section 3 we define the four period model and derive an expression for the MSE score we use. In section 4 we discuss how the weights in the model can be calculated. We then move on to applying various models to the hurricane data in sections 5, 6 and 7. Finally in section 8 we discuss our results.

2 Data

We now describe the hurricane data we will use for this study. The basic data set is HURDAT (Jarvinen et al., 1984). We extract the number of hurricanes making landfall each year, in each category, counting only the most intense landfall of storms that make multiple landfalls (this is the SSS variable in HURDAT). Table 1 shows the number of hurricanes in each of our four periods, for all hurricanes, and for increasingly intense hurricanes. Table 2 shows the same data, but now for numbers of hurricanes per year.

1	2	3	4	5	6	7
	00-42	43-64	65-94	95-05	00-42+65-94	43-64+95-05
1-5	76	43	38	24	101	68
2-5	46	27	21	17	67	44
3-5	27	18	14	10	41	28
4-5	8	5	3	1	11	6
5	1	0	2	0	3	0

Table 1: The columns show: (1) the categories of hurricanes being considered, (2) the number of land-falling hurricanes in these categories between 1900 and 1942, (3) the number of landfalling hurricanes in these categories between 1943 and 1964, (4) the number of landfalling hurricanes in these categories between 1965 and 1994, (5) the number of landfalling hurricanes in these categories between 1995 and 2005, (6) the number of landfalling hurricanes in these categories between 1900 and 1942 plus the number between 1965 and 1994, (7) the number of landfalling hurricanes in these categories between 1943 and 1964 plus the number between 1995 and 2005.

1	2	3	4	5	6	7
	00-42	43-64	65-94	95-05	00-42+65-94	43-64+95-05
1-5	1.77	1.95	1.27	2.27	1.38	2.06
2-5	1.07	1.23	0.70	1.55	0.92	1.33
3-5	0.63	0.82	0.47	0.91	0.56	0.85
4-5	0.19	0.23	0.10	0.09	0.15	0.18
5	0.02	0.00	0.07	0.00	0.04	0.00

Table 2: This table has the same contents as table 1, but the hurricane numbers are now given as average numbers per year, rather than total numbers.

3 The four period model

In this section we define the statistical model that represents hurricane activity as having been driven by four periods of differing underlying mean activity rates. In fact, we will consider the general case of k periods, and specialize to $k = 4$ later.

We will be taking an approach in which the historical data is considered as a single realisation of a random variable. All expectations are then with respect to this random variable. The random variable we use is:

$$Y_{i,j} \sim \text{Pois}(\lambda_i), (i = 1, 2, 3, k; j = 1, \dots, n_i) \quad (1)$$

$Y_{i,j}$ represents the hurricane counts during the j 'th year of the i 'th period. As mentioned above, we count our periods backwards in time, so that $j = 1$ corresponds to the most recent period, $j = 2$ the one before that, and so on.

Our predictions will be based on averages of the hurricane numbers during these k periods, which we write as:

$$\hat{\lambda}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{i,j} \quad (2)$$

We are trying to predict the hurricane count for next year, on the assumption that the hurricane activity next year is at the same level as the last few years. We are thus trying to predict the next value in period 1, which we write as Y_{1,n_1+1} .

Before considering weighted combinations of the hurricane number averages $\hat{\lambda}_i$ during these four periods, we first consider the two most obvious predictions for Y_{1,n_1+1} , which are the average of the hurricane counts in period 1:

$$\hat{Y}_{1,n_1+1} = \hat{\lambda}_1 \quad (3)$$

(we often call this the 'short baseline predictor')

and the average of the hurricane counts over all four periods:

$$Y_{1,n_1+1}^\dagger = \sum_{i=1}^k \frac{n_i}{n} \hat{\lambda}_i \quad (4)$$

where $n = \sum_{i=1}^k n_i$

(we often call this the 'long baseline predictor').

What are the properties of these two basic predictors?

3.1 Properties of the basic predictors

Rewriting the two basic predictors in terms of the real (unknown) hurricane rates in the k periods, plus (also unknown) noise, gives:

$$\hat{Y}_{1,n_1+1} = \lambda_1 + \frac{1}{n_1} \sum_{j=1}^{n_1} \varepsilon_{1,j} \quad (5)$$

$$Y_{1,n_1+1}^\dagger = \frac{n_1}{n} \left(\lambda_1 + \frac{1}{n_1} \sum_{j=1}^{n_1} \varepsilon_{1,j} \right) + \frac{n_2}{n} \left(\lambda_2 + \frac{1}{n_2} \sum_{j=1}^{n_2} \varepsilon_{2,j} \right) \quad (6)$$

$$+ \dots + \frac{n_k}{n} \left(\lambda_k + \frac{1}{n_k} \sum_{j=1}^{n_k} \varepsilon_{k,j} \right) \quad (7)$$

The prediction errors for these basic predictions are then:

$$Y_{1,n_1+1} - \hat{Y}_{1,n_1+1} = \lambda_1 + \varepsilon_{1,n_1+1} - \left(\lambda_1 + \frac{1}{n_1} \sum_{j=1}^{n_1} \varepsilon_{1,j} \right) \quad (8)$$

$$Y_{1,n_1+1} - Y_{1,n_1+1}^\dagger = \lambda_1 + \varepsilon_{1,n_1+1} - \frac{n_1}{n} \left(\lambda_1 + \frac{1}{n_1} \sum_{j=1}^{n_1} \varepsilon_{1,j} \right) - \frac{n_2}{n} \left(\lambda_2 + \frac{1}{n_2} \sum_{j=1}^{n_1} \varepsilon_{2,j} \right) \quad (9)$$

$$- \dots - \frac{n_k}{n} \left(\lambda_k + \frac{1}{n_k} \sum_{j=1}^{n_1} \varepsilon_{k,j} \right) \quad (10)$$

$$= \frac{1}{n} [(n - n_1)\lambda_1 - n_2\lambda_2 - \dots - n_k\lambda_k] \quad (11)$$

$$+ \varepsilon_{1,n_1+n} - \frac{1}{n} \left(\sum_{j=1}^{n_1} \varepsilon_{1,j} + \sum_{j=1}^{n_2} \varepsilon_{2,j} + \dots + \sum_{j=1}^{n_k} \varepsilon_{k,j} \right) \quad (12)$$

and the statistical properties of these predictions are:

$$E(Y_{1,n_1+1} - \hat{Y}_{1,n_1+1}) = 0 \quad (13)$$

$$E(E(Y_{1,n_1+1}) - \hat{Y}_{1,n_1+1}) = 0 \quad (14)$$

$$E(Y_{1,n_1+1} - Y_{1,n_1+1}^\dagger) = \frac{1}{n} [(n - n_1)\lambda_1 - n_2\lambda_2 - \dots - n_k\lambda_k] \quad (15)$$

$$E(E(Y_{1,n_1+1}) - Y_{1,n_1+1}^\dagger) = \frac{1}{n} [(n - n_1)\lambda_1 - n_2\lambda_2 - \dots - n_k\lambda_k] \quad (16)$$

$$\text{Var}(Y_{1,n_1+1} - \hat{Y}_{1,n_1+1}) = \left(1 + \frac{1}{n_1}\right) \lambda_1 \quad (17)$$

$$\text{Var}(E(Y_{1,n_1+1}) - \hat{Y}_{1,n_1+1}) = \frac{1}{n_1} \lambda_1 \quad (18)$$

$$\text{Var}(Y_{1,n_1+1} - Y_{1,n_1+1}^\dagger) = \frac{1}{n^2} [(n^2 + n_1)\lambda_1 + n_2\lambda_2 + \dots + n_k\lambda_k] \quad (19)$$

$$\text{Var}(E(Y_{1,n_1+1}) - Y_{1,n_1+1}^\dagger) = \frac{1}{n^2} [n_1\lambda_1 + n_2\lambda_2 + \dots + n_k\lambda_k] \quad (20)$$

and

$$\text{MSE}_1(\hat{Y}_{1,n_1+1}) = \left(1 + \frac{1}{n_1}\right) \lambda_1 \quad (21)$$

$$\text{MSE}_2(\hat{Y}_{1,n_1+1}) = \frac{1}{n_1} \lambda_1 \quad (22)$$

$$\begin{aligned} \text{MSE}_1(Y_{1,n_1+1}^\dagger) &= \frac{1}{n^2} [(n^2 - n_1)\lambda_1 - n_2\lambda_2 - \dots - n_k\lambda_k]^2 \\ &\quad + \frac{1}{n^2} [(n^2 + n_1)\lambda_1 + n_2\lambda_2 + \dots + n_k\lambda_k] \end{aligned} \quad (23)$$

$$\text{MSE}_2(Y_{1,n_1+1}^\dagger) = \frac{1}{n^2} [n_1\lambda_1 - n_2\lambda_2 - \dots - n_k\lambda_k]^2 \quad (24)$$

$$+ \frac{1}{n^2} [n_1\lambda_1 + n_2\lambda_2 + \dots + n_k\lambda_k] \quad (25)$$

One can ask the question: if we want to minimise the MSE, is it better to use the short baseline or the long baseline? The answer to this question can be obtained by the following comparison:

$$\text{MSE}(Y_{1,n_1+1}^\dagger) < \text{MSE}(\hat{Y}_{1,n_1+1}) \quad (26)$$

$$(27)$$

Neither side can be evaluated exactly in practice since they both depend on the unknown parameters $\lambda_1, \dots, \lambda_k$. But as a practical approach one could use plug-in estimators for the λ 's.

3.2 The mixed baseline model

We now move on to consider the 'mixed baseline model', in which our prediction consists of a weighted average of the mean hurricane numbers in the k periods, which we write as:

$$Y_{1,n_1+1}^*(\alpha) = \alpha_1 \hat{\lambda}_1 + \alpha_2 \hat{\lambda}_2 + \dots + \alpha_k \hat{\lambda}_k \quad (28)$$

subject to the constraint that the weights add up to 1, i.e.

$$\alpha_1 + \alpha_2 + \dots + \alpha_k = 1 \quad (29)$$

We now rewrite the prediction in terms of the real unknown levels of activity, plus noise:

$$Y_{1,n_1+1}^*(\alpha) = \alpha_1 \left(\lambda_1 + \frac{1}{n_1} \sum_{j=1}^{n_1} \varepsilon_{1,j} \right) + \alpha_2 \left(\lambda_2 + \frac{1}{n_2} \sum_{j=1}^{n_2} \varepsilon_{2,j} \right) + \dots + \alpha_k \left(\lambda_k + \frac{1}{n_k} \sum_{j=1}^{n_k} \varepsilon_{k,j} \right)$$

The prediction error is then

$$Y_{1,n_1+1} - Y_{1,n_1+1}^*(\alpha) = (1 - \alpha_1)\lambda_1 - \alpha_2\lambda_2 - \dots - \alpha_k\lambda_k \quad (30)$$

$$+ \varepsilon_{1,n_1+1} - \left(\frac{\alpha_1}{n_1} \sum_{j=1}^{n_1} \varepsilon_{1,j} + \frac{\alpha_2}{n_2} \sum_{j=1}^{n_2} \varepsilon_{2,j} + \dots + \frac{\alpha_k}{n_k} \sum_{j=1}^{n_k} \varepsilon_{k,j} \right) \quad (31)$$

The statistical properties of this prediction are:

$$E(Y_{1,n_1+1} - Y_{1,n_1+1}^*(\alpha)) = (1 - \alpha_1)\lambda_1 - \alpha_2\lambda_2 - \dots - \alpha_k\lambda_k \quad (32)$$

$$E(E(Y_{1,n_1+1}) - Y_{1,n_1+1}^*(\alpha)) = (1 - \alpha_1)\lambda_1 - \alpha_2\lambda_2 - \dots - \alpha_k\lambda_k \quad (33)$$

$$\text{Var}(Y_{1,n_1+1} - Y_{1,n_1+1}^*(\alpha)) = \left(1 + \frac{\alpha_1^2}{n_1} \right) \lambda_1 + \frac{\alpha_2^2}{n_2} \lambda_2 + \dots + \frac{\alpha_k^2}{n_k} \lambda_k \quad (34)$$

$$\text{Var}(E(Y_{1,n_1+1}) - Y_{1,n_1+1}^*(\alpha)) = \frac{\alpha_1^2}{n_1} \lambda_1 + \frac{\alpha_2^2}{n_2} \lambda_2 + \dots + \frac{\alpha_k^2}{n_k} \lambda_k \quad (35)$$

and

$$\text{MSE}_1(Y_{1,n_1+1}^*) = [(1 - \alpha_1)\lambda_1 - \alpha_2\lambda_2 - \dots - \alpha_k\lambda_k]^2 \quad (36)$$

$$+ \left[\left(1 + \frac{\alpha_1^2}{n_1} \right) \lambda_1 + \frac{\alpha_2^2}{n_2} \lambda_2 + \dots + \frac{\alpha_k^2}{n_k} \lambda_k \right] \quad (37)$$

$$\text{MSE}_2(Y_{1,n_1+1}^*) = [(1 - \alpha_1)\lambda_1 - \alpha_2\lambda_2 - \dots - \alpha_k\lambda_k]^2 \quad (38)$$

$$+ \left[\frac{\alpha_1^2}{n_1} \lambda_1 + \frac{\alpha_2^2}{n_2} \lambda_2 + \dots + \frac{\alpha_k^2}{n_k} \lambda_k \right] \quad (39)$$

Our goal is to find the values of the α 's that minimize this MSE subject to

$$\alpha_1 + \alpha_2 + \dots + \alpha_k = 1 \quad (40)$$

$$\alpha_i \geq 0 \quad (41)$$

for $i = (1, 2, \dots, k)$.

3.3 Note on definition of the model

We have defined our model in terms of weights on averages over periods of years. We could also have defined it in terms of weights on individual years, with these annual weights being constant in each of the k periods. Since:

$$Y_{1,n_1+1}^*(\alpha) = \alpha_1 \hat{\lambda}_1 + \alpha_2 \hat{\lambda}_2 + \dots + \alpha_k \hat{\lambda}_k \quad (42)$$

$$= \sum_{i=1}^k \alpha_i \hat{\lambda}_i \quad (43)$$

$$= \sum_{i=1}^k \alpha_i \left(\frac{1}{n_i} \sum_{j=1}^{n_i} Y_{i,j} \right) \quad (44)$$

$$= \sum_{i=1}^k \left(\sum_{j=1}^{n_i} \frac{\alpha_i}{n_i} Y_{i,j} \right) \quad (45)$$

We see that the annual weight on year j in period i is given by $\frac{\alpha_i}{n_i}$.

These annual weights give us a useful alternative interpretation of the weights that come out of the model when applied to real data, as we will see below.

4 Calculating the optimal weights

For the model described above in section 3, how should we estimate the optimal weights? Ideally we would derive exact expressions for the weights, that extend the expressions given for the $k = 2$ case in Jewson et al. (2005). A derivation for exact expressions for the weights, taking into account the constraint that the weights should sum to one, but not the constraint that the weights should be positive, is given in appendix 1. However, when we tested the resulting expressions on real data we found that they gave negative weights, which are not acceptable.

As a simple alternative for the $k = 4$ case we are interested in, we then resorted to a brute-force search through all possibilities. If we assume that we only need accuracy to two decimal places then each weight can only have values from 0.00 to 1.00, which is 101 possibilities. The fourth weight is given by the other three, and there are thus only $101^3 = 1,030,301$ parameter sets to choose from. We have tested this method using Fortran on a desktop computer, and completing the search for the weights that give the lowest RMSE took only a few seconds.

There are other (more efficient) numerical methods that one might consider, especially if k were large, but brute-force searching turned out to be perfectly adequate for our needs in this particular case.

5 Results for all hurricanes

We now show some results for a number of models applied to the data for all hurricanes, shown in table 3. Values of bias, standard deviation of errors and RMSE are all calculated using the four level model, to make them consistent. The RMSE's are shown as what we call 'RMSE2', which is the RMSE on the prediction of the rate rather than the number of hurricanes (aka the standard error on the rate). The weights are shown in chronological order, and are shown both as the original weights used in the model and the annual weights discussed above in section 3.3. The annual weights are multiplied by 106 to make them easier to understand.

5.0.1 Model 1: Long term baseline

The first model we consider is the 'long term baseline', which estimates the future rates as an average of hurricane numbers during the period 1900-2005. This model gives a prediction of 1.72 hurricanes per year. This prediction has a low standard deviation (of 0.13), because it is based on many years of data, but has a significant bias (of 0.56). The bias arises because of the assumptions in our four period model, which are that the future will come from the same climate as the last 11 years, and the last 11 years have seen more hurricanes than the 106 years baseline. The large bias leads to a large RMSE.

5.0.2 Model 2: Short term baseline

The second simple model we consider is the ‘short term baseline’, which estimates the future rates as an average of hurricane numbers during the period 1995-2005. This model gives a much higher prediction of future hurricane rates, of 2.27 hurricanes per year. Given our assumptions, this model has no bias. But it does have a large standard deviation of errors, because of the short period of data used. The resulting RMSE is lower than for the long term baseline model, but is still rather large (0.45).

The subsequent models are all an attempt to give a better prediction (i.e. a prediction with lower RMSE) than either of these two very simple models.

5.0.3 Model 3: Optimal combination of the long and short baselines

Our next model considers the optimal combination of the long and short baselines, just as in Jewson et al. (2005), but now with the weights calculated to minimise the RMSE as given by equation 38 above (i.e. to minimise the RMSE calculated in the context of the four level model). The optimal weights come out to be 66% on the recent data and 34% on the earlier data. The annual weights show this to be equivalent to putting just over 16 times as much weight on the values from the recent years than the earlier years. The RMSE in this model cannot be worse than the two models that go into it, by definition. In fact, it is somewhat better than the short baseline model.

One note: it is important to realise that this model does not assume that there is a 66% chance that the next year will be at the same level as the last 11 years, and a 34% chance that the next year will revert to the long term baseline. It assumes that there is a 100% chance that the next year will be at the same level as the last 11 years. The weighting comes in as a mathematical result of trying to reduce the errors in our prediction. If one wanted to include some probability of reverting to a lower level in the future, one would need to do that some other way.

5.0.4 Model 4: Average of the active periods

As mentioned in the introduction, no-one really believes that the period 1900-1994 is uniform and without fluctuations in the hurricane rate. Model 4 is the first to try and take that into account. It does so by ignoring the two periods with lower rates, and taking a straight average of the two periods with a higher rate. This model performs well in terms of RMSE, and is the best model so far.

5.0.5 Model 5: Optimal combination of the active periods

However, the earlier active period was not as active as the recent active period, and so perhaps we should put more weight on the recent period than comes from the straight average of all active years that we used in model 4. Model 5 tries an optimal combination of the two active periods to capture this. We get a lower RMSE than model 4, as we must by construction (because we’ve added an extra parameter) and slightly higher weights on the recent data. The forecast is slightly higher than model 4 because of these higher weights on the recent (and more active) period.

5.0.6 Model 6: One active period, one inactive period

A potential shortcoming of models 4 and 5 is that they don’t use the whole available data set. We now attempt ways to use the whole data set, but still resolve the idea that there was an active period in the middle of the last century. Model 6 does this in the simplest way possible by aggregating the two less active phases into one period, and the two active phases into another period. It then combines the average levels across these two aggregated periods with weights that minimise RMSE. The resulting weights are very interesting: the algorithm that determines the optimal weights actually sets the weights on the aggregated inactive period to zero, and model 6 thus becomes equivalent to a straight average of the two active periods. This is telling us that there is *not* a potential benefit to be had from using the data from the inactive periods: although using such data would reduce the contribution to the RMSE from the standard deviation of the errors, it would increase the contribution to the RMSE from the bias by more.

5.0.7 Model 7: Two active periods, one inactive period

An extension of model 6 above is to distinguish between the rates in the two active periods, but not in the two inactive periods. As might be expected given the results for model 6, this method again drops the inactive data from the combination, and thus reverts to model 5.

5.0.8 Model 8: four distinct periods

The most complex model we consider takes four periods at different levels and combines them using four independent weights (subject to the constraints). By definition this model must give the best RMSE score, and it does. However, we suspect that the score may not be not *statistically significantly* lower than the scores from models 5 and 7, and the forecast is very similar. Model 8 gives a zero weight to the data from period 2, and only a very small weight to the data from the earlier inactive period.

5.1 Discussion

Consideration of these models has yielded the interesting result that the data from the two inactive periods cannot be used in a very useful way to improve our forecast, because the rates during the inactive periods are too different. The data from the earlier active period, however, can definitely be used. Our best model in terms of RMSE (model 8) does put a small weight on the data from the earlier inactive period. However, the RMSE benefit is very small, and this model has 3 parameters versus only 1 parameter for model 5. If we did an analysis of whether the benefit from using model 8 over model 5 were statistically significant, we suspect it wouldn't be.

6 Results for intense hurricanes

The results for intense hurricanes (shown in table 4) are broadly similar, but with some specific differences. We only comment on the differences.

6.0.1 Model 2: short term baseline

It is interesting to note that the short term baseline model now does worse than the long term baseline model, which is a reverse of the results for total hurricane numbers. This is because there are fewer cat 3-5 storms, and using only 11 years of data works less well than it does for cat 1-5 storms.

6.0.2 Model 5: optimal combination of the two active periods

The optimal combination of the two active periods is now more or less the same as the straight average. This is because the optimal combination puts more weight on the earlier period than it does for cat 1-5 hurricanes. This is again because there are fewer cat 3-5 storms, so we need to use more years of data to estimate them accurately, even if that data is not ideal.

7 Results for basin hurricanes

For future reference, we also include results for basin hurricanes, in tables 5 and 6.

8 Conclusions

We have considered the question of how to combine historical hurricane numbers to make a prediction of the future. We have made this question mathematically tractable by assuming that the hurricane rates of the last 106 years were characterized by four periods of differing levels of activity, and that future activity will continue at the current level. We then look at different ways that one might combine estimates of activity from these four periods.

There are various conclusions we come to. Firstly, using data from the two inactive periods doesn't add much, if anything, to our predictions. This is because the levels of activity during the inactive periods were low relative to the current level of activity. Any potential benefit from using more data is wiped out by the fact that the data is at the wrong level. Secondly, using data from the earlier inactive period can definitely help our predictions. This data is not at exactly the same level as the current period of high activity, but is sufficiently close that the bias introduced by using this data is small compared with the beneficial reduction in variance. Overall, using this data reduces the RMSE of our predictions considerably.

The biggest flaw in this study is that we don't have an objective method for choosing the best model from the 8 models we have considered. The 3 parameter model (model 8) gives the best results (by construction), but may be overfitted relative to simpler models such as model 4 or model 5. Our next goal is therefore to consider how to select between these models.

A Exact derivation for weights for arbitrary k

We derive exact expressions for the weights as follows. We define the problem as needing to minimise a cost function L , given below. This cost function includes a Lagrange multiplier to account for the constraint that the weights must sum to 1.

$$L = \left[[(1 - \alpha_1)\lambda_1 - \alpha_2\lambda_2 - \dots - \alpha_k\lambda_k]^2 + \left[\left(1 + \frac{\alpha_1^2}{n_1}\right)\lambda_1 + \frac{\alpha_2^2}{n_2}\lambda_2 + \dots + \frac{\alpha_k^2}{n_k}\lambda_k \right] - \gamma \left(\sum_{i=1}^k \alpha_i - 1 \right) \right] \quad (46)$$

We then differentiate this cost function by the weight vector α and the lagrange multiplier γ , giving a $k + 1$ by 1 matrix of derivatives:

$$\begin{pmatrix} \frac{\partial}{\partial \alpha} \text{MSE}[\hat{Y}_{1,n_1+1}^*] \\ \frac{\partial}{\partial \gamma} \text{MSE}[\hat{Y}_{1,n_1+1}^*] \end{pmatrix} = \begin{pmatrix} -2\lambda_1[(1 - \alpha_1)\lambda_1 - \lambda_2\alpha_2 - \dots - \lambda_k\alpha_k] + 2\lambda_1\frac{\alpha_1}{n_1} - \gamma \\ -2\lambda_2[(1 - \alpha_1)\lambda_1 - \lambda_2\alpha_2 - \dots - \lambda_k\alpha_k] + 2\lambda_1\frac{\alpha_2}{n_2} - \gamma \\ \vdots \\ -2\lambda_k[(1 - \alpha_1)\lambda_1 - \lambda_2\alpha_2 - \dots - \lambda_k\alpha_k] + 2\lambda_1\frac{\alpha_k}{n_k} - \gamma \\ -\alpha_1 - \alpha_2 - \dots - \alpha_k + 1 \end{pmatrix} \quad (47)$$

To find the minimum, we set:

$$\begin{pmatrix} \frac{\partial}{\partial \alpha} \text{MSE}[\hat{Y}_{1,n_1+1}^*] \\ \frac{\partial}{\partial \gamma} \text{MSE}[\hat{Y}_{1,n_1+1}^*] \end{pmatrix} = 0 \quad (48)$$

and rearranging the equations we obtain

$$-2\lambda_1\lambda_1 + 2\left(\lambda_1\lambda_1 + \lambda_1\frac{1}{n_1}\right)\alpha_1 + 2\lambda_1\lambda_2\alpha_2 + \dots + 2\lambda_1\lambda_k\alpha_k - \gamma = 0 \quad (49)$$

$$-2\lambda_1\lambda_2 + 2\lambda_1\lambda_2\alpha_1 + 2\left(\lambda_2\lambda_2 + \lambda_1\frac{1}{n_2}\right)\alpha_2 + \dots + 2\lambda_2\lambda_k\alpha_k - \gamma = 0 \quad (50)$$

$$\vdots \quad (51)$$

$$-2\lambda_1\lambda_k + 2\lambda_1\lambda_k\alpha_1 + 2\lambda_k\lambda_2\alpha_2 + \dots + 2\left(\lambda_k\lambda_k + \lambda_1\frac{1}{n_k}\right)\alpha_k - \gamma = 0 \quad (52)$$

$$1 - \alpha_1 - \alpha_2 - \dots - \alpha_k - 0 \cdot \gamma = 0 \quad (53)$$

This set of equations can be rewritten in matrix form as

$$\begin{pmatrix} 2\left(\lambda_1\lambda_1 + \lambda_1\frac{1}{n_1}\right) & 2\lambda_1\lambda_2 & \dots & 2\lambda_1\lambda_k & -1 \\ 2\lambda_1\lambda_2 & 2\left(\lambda_2\lambda_2 + \lambda_1\frac{1}{n_2}\right) & \dots & 2\lambda_2\lambda_k & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 2\lambda_1\lambda_k & 2\lambda_2\lambda_k & \dots & 2\left(\lambda_k\lambda_k + \lambda_1\frac{1}{n_k}\right) & -1 \\ -1 & -1 & \dots & -1 & 0 \end{pmatrix} \times \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \\ \gamma \end{pmatrix} = \begin{pmatrix} 2\lambda_1\lambda_1 \\ 2\lambda_1\lambda_2 \\ \vdots \\ 2\lambda_1\lambda_k \\ -1 \end{pmatrix} \quad (54)$$

In shorthand notation we can write

$$M \times \alpha = c \quad (55)$$

where M is a symmetric $(k + 1) \times (k + 1)$ matrix of coefficients, α is $k + 1$ column vector of weights α_i and Lagrange multiplier γ and c is a $k + 1$ column vector of constants.

The solution for alpha is then obtained by

$$\alpha = M^{-1} \times c \quad (56)$$

A.1 Special cases

The simplest interesting case is when $k = 2$, which is equivalent to the original model of Jewson et al. (2005). We now show that the above expressions agree with that model for this special case.

Setting $k = 2$, we get

$$\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \gamma \end{pmatrix} \quad (57)$$

$$M = \begin{pmatrix} 2\left(\lambda_1\lambda_1 + \lambda_1\frac{1}{n_1}\right) & 2\lambda_1\lambda_2 & -1 \\ 2\lambda_2\lambda_1 & 2\left(\lambda_2\lambda_2 + \lambda_1\frac{1}{n_2}\right) & -1 \\ -1 & -1 & 0 \end{pmatrix} \quad (58)$$

$$c = \begin{pmatrix} 2\lambda_1\lambda_1 \\ 2\lambda_1\lambda_2 \\ -1 \end{pmatrix} \quad (59)$$

and so

$$M^{-1} = \begin{pmatrix} \frac{1}{2\left[\frac{\lambda_1}{n_1} + (\lambda_1 - \lambda_2)^2\right]} & -\frac{1}{2\left[\frac{\lambda_1}{n_1} + (\lambda_1 - \lambda_2)^2\right]} & -\frac{\frac{\lambda_2}{n_2} - \lambda_1\lambda_2 + \lambda_2^2}{\frac{\lambda_1}{n_1} + (\lambda_1 - \lambda_2)^2} \\ -\frac{1}{2\left[\frac{\lambda_1}{n_1} + (\lambda_1 - \lambda_2)^2\right]} & \frac{1}{2\left[\frac{\lambda_1}{n_1} + (\lambda_1 - \lambda_2)^2\right]} & -\frac{\frac{\lambda_1}{n_1} - \lambda_1 + \lambda_1^2}{\frac{\lambda_1}{n_1} + (\lambda_1 - \lambda_2)^2} \\ -\frac{\frac{\lambda_2}{n_2} - \lambda_1\lambda_2 + \lambda_2^2}{\frac{\lambda_1}{n_1} + (\lambda_1 - \lambda_2)^2} & -\frac{\frac{\lambda_1}{n_1} - \lambda_1\lambda_2 + \lambda_1^2}{\frac{\lambda_1}{n_1} + (\lambda_1 - \lambda_2)^2} & -2\frac{\frac{\lambda_1\lambda_2}{n_1n_2} + \frac{\lambda_1^2\lambda_2}{n_2} + \frac{\lambda_1\lambda_2^2}{n_1}}{\frac{\lambda_1}{n_1} + (\lambda_1 - \lambda_2)^2} \end{pmatrix} \quad (60)$$

and

$$\alpha = \frac{1}{n_1n_2(\lambda_1 - \lambda_2)^2 + n_2\lambda_1 + n_1\lambda_2} \begin{pmatrix} n_1n_2(\lambda_1 - \lambda_2)^2 + n_1\lambda_2 \\ n_2\lambda_1 \\ 2[\lambda_1\lambda_2 + n_2(\lambda_1\lambda_2^2 - \lambda_1^2\lambda_2)] \end{pmatrix} \quad (61)$$

The first two weights in this expression agree with the corresponding expression in Jewson et al. (2005).

References

- J Elsner, T Jagger, and X Niu. Changes in the rates of North Atlantic major hurricane activity during the 20th Century. *Geophysical Research Letters*, 27:1743–1746, 2000.
- J Elsner and C Schmertmann. Improving extended-range seasonal predictions of intense Atlantic hurricane activity. *Weather and Forecasting*, 3:345–351, 1993.
- W Gray. Predicting Atlantic basin seasonal tropical cyclone activity by 1 August. *Weather and Forecasting*, 8:73–86, 1993.
- W Gray. Predicting Atlantic basin seasonal tropical cyclone activity by 1 June. *Weather and Forecasting*, 9:103–115, 1994.
- W Gray, C W Landsea, P W Mielke, and K J Berry. Predicting Atlantic seasonal hurricane activity 6-11 months in advance. *Weather and Forecasting*, 7:440–455, 1992.
- W M Gray. Atlantic seasonal hurricane frequency - Part II: Forecasting its variability. *Monthly Weather Review*, 112:1669–1683, 1984.

- B Jarvinen, C Neumann, and M Davis. A tropical cyclone data tape for the North Atlantic Basin, 1886-1983: Contents, limitations, and uses. Technical report, NOAA Technical Memorandum NWS NHC 22, 1984.
- S Jewson, C Casey, and J Penzer. Year ahead prediction of US landfalling hurricane numbers: the optimal combination of long and short baselines. *arxiv:physics/0512113*, 2005.
- C W Landsea. *El Nino-Southern Oscillation and the seasonal predictability of tropical cyclones - Multiscale variability and global and regional impacts*, in: *El Nino - Implications of Multiscale Variability on Natural Ecosystems and Society*. editors: H F Diaz and V Markgraf, 2000.
- C W Landsea, G D Bell, and W M Gray. The extremely active 1995 atlantic hurricane season: Environmental conditions and verification of seasonal forecasts. *Monthly Weather Review*, 126:1174–1193, 1999.
- G S Lehmiller, T B Kimberlain, and J B Elsner. Seasonal prediction models for North Atlantic basic hurricane location. *Monthly Weather Review*, 125:1780–1791, 1997.
- M Saunders and A Lea. Seasonal prediction of hurricane activity reaching the coast of the United States. *Nature*, 434:1005–1008, 2005.

1	2	3	4	5	6	7	8
Model	Model Name(No. Yrs)	Fcst	RMSE2	Pctage	Bias,SD	Weights (a4,a3,a2,a1)	Scaled Annual Weights
1	long-term BL (106)	1.72	0.570	33.21	0.56, 0.13	0.41, 0.21, 0.28, 0.10	1.00, 1.00, 1.00, 1.00
2	short-term BL (11)	2.27	0.455	20.00	0.00, 0.45	0.00, 0.00, 0.00, 1.00	0.00, 0.00, 0.00, 9.64
3	2 pds, pre/post 1994/5 (106)	2.06	0.369	17.92	0.21, 0.30	0.15, 0.08, 0.11, 0.66	0.38, 0.38, 0.38, 6.36
4	2 active pds (33)	2.06	0.328	15.91	0.21, 0.25	0.00, 0.67, 0.00, 0.33	0.00, 3.21, 0.00, 3.21
5	2 active pds, optimal (33)	2.11	0.315	14.93	0.17, 0.27	0.00, 0.52, 0.00, 0.48	0.00, 2.51, 0.00, 4.62
6	2 periods,1-Act/1-Inact (106)	2.06	0.328	15.91	0.21, 0.25	0.00, 0.67, 0.00, 0.33	0.00, 3.21, 0.00, 3.21
7	3 pds,2-Act/1-Inact (106)	2.11	0.315	14.93	0.17, 0.27	0.00, 0.52, 0.00, 0.48	0.00, 2.51, 0.00, 4.62
8	4 pds (106)	2.09	0.312	14.97	0.19, 0.25	0.09, 0.43, 0.00, 0.47	0.23, 2.09, 0.00, 4.55

Table 3: Mixed baseline results for cat 1-5 US landfalling hurricanes

1	2	3	4	5	6	7	8
Model	Model Name(No. Yrs)	Fcst	RMSE2	Pctage	Bias,SD	Weights (a4,a3,a2,a1)	Scaled Annual Weights
1	long-term BL (106)	0.65	0.270	41.44	0.26, 0.08	0.41, 0.21, 0.28, 0.10	1.00, 1.00, 1.00, 1.00
2	short-term BL (11)	0.91	0.287	31.62	0.00, 0.29	0.00, 0.00, 0.00, 1.00	0.00, 0.00, 0.00, 9.64
3	2 pds, pre/post 1994/5 (106)	0.77	0.207	26.89	0.14, 0.15	0.22, 0.11, 0.15, 0.52	0.54, 0.54, 0.54, 5.01
4	2 active pds (33)	0.85	0.171	20.20	0.06, 0.16	0.00, 0.67, 0.00, 0.33	0.00, 3.21, 0.00, 3.21
5	2 active pds, optimal (33)	0.85	0.171	20.14	0.06, 0.16	0.00, 0.64, 0.00, 0.36	0.00, 3.11, 0.00, 3.42
6	2 periods,1-Act/1-Inact (106)	0.83	0.170	20.49	0.08, 0.15	0.04, 0.62, 0.03, 0.31	0.10, 2.98, 0.10, 2.98
7	3 pds,2-Act/1-Inact (106)	0.83	0.169	20.41	0.08, 0.15	0.05, 0.57, 0.03, 0.35	0.12, 2.76, 0.12, 3.33
8	4 pds (106)	0.82	0.165	20.19	0.09, 0.14	0.15, 0.52, 0.00, 0.33	0.37, 2.49, 0.00, 3.20

Table 4: Mixed baseline results for cat 3-5 US landfalling hurricanes

1	2	3	4	5	6	7	8
Model	Model Name(No. Yrs)	Fcst	RMSE2	Pctage	Bias,SD	Weights (a4,a3,a2,a1)	Scaled Annual Weights
1	long-term BL (106)	5.28	3.179	60.18	3.17, 0.22	0.41, 0.21, 0.28, 0.10	1.00, 1.00, 1.00, 1.00
2	short-term BL (11)	8.45	0.877	10.37	0.00, 0.88	0.00, 0.00, 0.00, 1.00	0.00, 0.00, 0.00, 9.64
3	2 pds, pre/post 1994/5 (106)	8.25	0.851	10.32	0.21, 0.83	0.03, 0.01, 0.02, 0.94	0.06, 0.06, 0.06, 9.08
4	2 active pds (33)	6.91	1.612	23.33	1.55, 0.46	0.00, 0.67, 0.00, 0.33	0.00, 3.21, 0.00, 3.21
5	2 active pds, optimal (33)	8.18	0.823	10.06	0.28, 0.77	0.00, 0.12, 0.00, 0.88	0.00, 0.58, 0.00, 8.48
6	2 periods,1-Act/1-Inact (106)	6.91	1.612	23.33	1.55, 0.46	0.00, 0.67, 0.00, 0.33	0.00, 3.21, 0.00, 3.21
7	3 pds,2-Act/1-Inact (106)	8.18	0.823	10.06	0.28, 0.77	0.00, 0.12, 0.00, 0.88	0.00, 0.58, 0.00, 8.48
8	4 pds (106)	8.18	0.823	10.06	0.28, 0.77	0.00, 0.12, 0.00, 0.88	0.00, 0.58, 0.00, 8.48

Table 5: Mixed baseline results for cat 1-5 Atlantic basin hurricanes

1	2	3	4	5	6	7	8
Model	Model Name(No. Yrs)	Fcst	RMSE2	Pctage	Bias,SD	Weights (a4,a3,a2,a1)	Scaled Annual Weights
1	long-term BL (106)	2.17	1.926	88.78	1.92, 0.14	0.41, 0.21, 0.28, 0.10	1.00, 1.00, 1.00, 1.00
2	short-term BL (11)	4.09	0.610	14.91	0.00, 0.61	0.00, 0.00, 0.00, 1.00	0.00, 0.00, 0.00, 9.64
3	2 pds, pre/post 1994/5 (106)	3.93	0.587	14.93	0.16, 0.56	0.03, 0.02, 0.02, 0.93	0.08, 0.08, 0.08, 8.91
4	2 active pds (33)	3.64	0.563	15.48	0.45, 0.33	0.00, 0.67, 0.00, 0.33	0.00, 3.21, 0.00, 3.21
5	2 active pds, optimal (33)	3.84	0.482	12.57	0.26, 0.41	0.00, 0.38, 0.00, 0.63	0.00, 1.81, 0.00, 6.02
6	2 periods,1-Act/1-Inact (106)	3.64	0.563	15.48	0.45, 0.33	0.00, 0.67, 0.00, 0.33	0.00, 3.21, 0.00, 3.21
7	3 pds,2-Act/1-Inact (106)	3.84	0.482	12.57	0.26, 0.41	0.00, 0.38, 0.00, 0.63	0.00, 1.81, 0.00, 6.02
8	4 pds (106)	3.83	0.482	12.57	0.26, 0.41	0.00, 0.38, 0.00, 0.62	0.00, 1.81, 0.00, 6.02

Table 6: Mixed baseline results for cat 3-5 Atlantic basin hurricanes